$$= \frac{n}{2} + \frac{1}{2} \ln(2\pi)^n |K| \tag{9.43}$$

$$= \frac{1}{2} \ln(2\pi e)^n |K| \text{ nats} \tag{9.44}$$

$$= \frac{1}{2} \log(2\pi e)^n |K| \text{ bits}. \quad \Box \tag{9.45}$$

9.5 RELATIVE ENTROPY AND MUTUAL INFORMATION

We now extend the definition of two familiar quantities, $D(f\|g)$ and $I(X;Y)$ to probability densities.

Definition: The *relative entropy* (or *Kullback Leibler distance*) $D(f\|g)$ between two densities f and g is defined by

$$D(f\|g) = \int f \log \frac{f}{g}. \tag{9.46}$$

Note that $D(f\|g)$ is finite only if the support set of f is contained in the support set of g. (Motivated by continuity, we set $0 \log \frac{0}{0} = 0$.)

Definition: The *mutual information* $I(X;Y)$ between two random variables with joint density $f(x,y)$ is defined as

$$I(X;Y) = \int f(x,y) \log \frac{f(x,y)}{f(x)f(y)}\, dx\, dy. \tag{9.47}$$

From the definition it is clear that

$$I(X;Y) = h(X) - h(X|Y) = h(Y) - h(Y|X) \tag{9.48}$$

and

$$I(X;Y) = D(f(x,y)\|f(x)f(y)). \tag{9.49}$$

The properties of $D(f\|g)$ and $I(X;Y)$ are the same as in the discrete case. In particular, the mutual information between two random variables is the limit of the mutual information between their quantized versions, since

$$I(X^\Delta; Y^\Delta) = H(X^\Delta) - H(X^\Delta|Y^\Delta) \tag{9.50}$$

$$\approx h(X) - \log\Delta - (h(X|Y) - \log\Delta) \tag{9.51}$$

$$= I(X;Y). \tag{9.52}$$